A december 16.-i gyakorlat témája

1.) Egy pénzdarabról ellenőrizni akarjuk, hogy igaz-e az a hipotézis, amely szerint ez az érme legalább $\frac{3}{4}$ valószínűséggel esik a fej és legfeljebb $\frac{1}{4}$ valószínűséggel az írás oldalára. Ennek érdekében feldobjuk a pénzdarabot 30 000 alkalommal, és a következő döntési szabályt hozzuk. Választunk egy k számot, és akkor fogadjuk el a hipotézist helyesnek, ha legalább k fejdobás történt. Legalább mekkorának kell válassztanunk ezt a k számot, ha azt akarjuk, hogy egy a hipotézist teljesítő pénzdarab esetén legalább 0.9 valószínűséggel döntsünk úgy, hogy a hipotézis teljesül?

Megoldás: Vezessük be a következő valószínűségi változókat: $\xi_j = 1$, ha a j-ik dobás eredménye fej, $\xi_j = 0$, ha a j-ik dobás eredménye írás, $1 \le j \le 30\,000$, $S = S_{30000} = \sum\limits_{j=1}^{30\,000} \xi_j$. Ha a fejdobás eredményének valószínűsége pontosan $\dfrac{3}{4}$, $E\xi_j = \dfrac{3}{4}$, $E\xi_j^2 = \dfrac{3}{4}$, $\operatorname{Var}\xi_j = E\xi_j^2 - (E\xi_j)^2 = \dfrac{3}{16}$, $ES = 30\,000 E\xi_j = 22\,500$, $\operatorname{Var} S = 30\,000 \operatorname{Var}\xi_j = 5625 = 75^2$. Innen és a centrális határeloszlástételből,

$$P(S > k) = P\left(\frac{S - ES}{\sqrt{\operatorname{Var} S}} > \frac{k - 22\,500}{75}\right) = 1 - P\left(\frac{S - ES}{\sqrt{\operatorname{Var} S}} \le \frac{k - 22\,500}{75}\right)$$
$$\sim 1 - \Phi\left(\frac{k - 22\,500}{75}\right).$$

Válasszuk a k számot úgy, hogy a fenti valószínűség körülbelül 0.9 legyen. Ekkor a $\Phi\left(\dfrac{k - 22\,500}{75}\right) = 0.1$ vagy ami ezzel ekvivalens, a $\Phi\left(\dfrac{22\,500 - k}{75}\right) = 0.9$ egyenletet kell kielégítenünk. A normális eloszlás-táblázat alapján $\dfrac{22\,500 - k}{75} \sim 1.28$, ami azt jelenti, hogy $k = 22\,500 - 75 \times 1.28$ és $p = \dfrac{3}{4}$ esetén annak valószínűsége, hogy a fejdobások száma nagyobb mint $k = 22\,500 - 75 \times 1.28 = 22\,212$ és $p = \dfrac{3}{4}$ esetében annak valószínűsége, hogy legalább ennyi fejdobás történik körülbelül 0.9. Ha $p \ge \dfrac{3}{4}$, akkor ez a valószínűség nagyobb. Ezért a $k = 22\,212$ helyes választás.

Megtárgyalunk egy fontos diszkrét eloszlást, amellyel időhiány miatt nem foglalkoztunk.

Poisson eloszlás definiciója. *Egy ξ valószínűségi változó λ paraméterű Poisson eloszlású, ha ξ nem negatív egész értékeket vesz fel, és*

$$P(\xi = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, 2, \dots$$

2.) Lássuk be, hogy az előző definicióban valóban egy eloszlást definiáltunk.

Megoldás: $P(\xi = k) = \frac{\lambda^k}{k!} e^{-\lambda} \geq 0$ minden $k = 0, 1, 2, \ldots$ számra. Azt kell még megmutatni, hogy $\sum_{k=0}^{\infty} P(\xi = k) = 1$. Viszont

$$\sum_{k=0}^{\infty} P(\xi = k) = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} e^{\lambda} = 1.$$

3.) Ha ξ és η két független, λ illetve μ paraméterű Poisson eloszlású valószínűségi változó, akkor $\xi + \eta$ $\lambda + \mu$ paraméterű Poisson eloszlású valószínűségi változó.
Megoldás:

$$\begin{aligned} P(\xi + \eta = k) &= \sum_{j=0}^{k} P(\xi = j, \eta = k - j) = \sum_{j=0}^{k} P(\xi = j) P(\eta = k - j) \\ &= \sum_{j=0}^{k} \frac{\lambda^j}{j!} e^{-\lambda} \frac{\mu^{k-j}}{(k-j)!} e^{-\mu} = \frac{e^{-(\lambda+\mu)}}{k!} \sum_{j=0}^{k} \frac{k!}{j!(k-j)!} \lambda^j \mu^{(k-j)} \\ &= \frac{e^{-(\lambda+\mu)}}{k!} (\lambda + \mu)^k. \end{aligned}$$

4.) Legyen ξ standard normális eloszlású valószínűségi változó, azaz legyen ξ sűrűségfüggvénye $f(x) = \dfrac{1}{\sqrt{2\pi}} e^{-x^2/2}$, és legyen t valós szám. Számoljuk ki az $e^{t\xi}$ valószínűségi változó $Ee^{t\xi}$ várható értékét.
Megoldás:

$$\begin{aligned} Ee^{t\xi} &= \int_{-\infty}^{\infty} e^{tu} \frac{1}{\sqrt{2\pi}} e^{-u^2/2}\, du = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{tu - u^2/2 - t^2/2 + t^2/2}\, du \\ &= e^{t^2/2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-(t-u)^2/2}\, du = e^{t^2/2}. \end{aligned}$$

DOLGOZAT FELADATOK

(A programozó matematikus gyakorlaton)

1.) Legyen egy urnában 10 piros és 40 fehér golyó. Húzzunk ki az urnából 10 golyót visszatevés nélkül. Mi a kihízottt piros golyók számának a várható értéke és szórásnégyzete?

2.) Legyen ξ exponenciális eloszlású valószínűségi változó λ paraméterrel, azaz legyen ξ eloszlásfüggvénye $F(x) = 1 - e^{-\lambda x}$, ha $x \geq 0$, $F(x) = 0$, ha $x < 0$. Számoljuk ki ξ^3 sűrűségfüggvényét.

3.) Legyen ξ egyenletes eloszlású valószínűségi változó a $[0, 1]$ intervallumban, azaz legyen sűrűségfüggvéye $f(x) = 1$, ha $0 \leq x \leq 1$, $f(x) = 0$, ha $x < 0$ vagy $x > 1$. Számítsuk ki a ξ^5 valószínűségi változó várható értékét és szórásnégyzetét.

4.) Legyen a ξ valószínűségi változó sűrűségfüggvénye $f(x) = \dfrac{1}{2\pi}\left(1+\dfrac{1}{2}\sin x\right)$, ha $0 \le x \le 2\pi$, $f(x)=0$, ha $x<0$ vagy $x>2\pi$. Számítsuk ki a ξ valószínűségi változó várható értékét és szórásnégyzetét.

5.) Egy szabályos dobókockát feldobunk 1200 alkalommal egymástól függetlenül, és összeadjuk a páros értékű dobások eredményét. Adjunk jó közelítő becslést a centrális határeloszlástétel és egy normális eloszlástáblázat segítségével arra, hogy ez az összeg 2280 és 2500 közé esik.

6.) Legyen adva n darab $\xi_1,\dots,\xi_n$ valószínűségi változó egy $(\Omega,\mathcal{A},P)$ valószínűségi mezőn. Mikor mondjuk, hogy ezek a $\xi_1,\dots,\xi_n$ valószínűségi változók függetlenek egymástól?

MEGOLDÁSOK

1.) Vezessük be a következő ξ_j, $1\le j\le 10$ valószínűségi változókat: $\xi_j=1$, ha a J-ik húzás piros $\xi_j=0$, ha a j-ik húzás fehér. Ekkor minket az $S=\sum\limits_{j=1}^{10}\xi_j$ valószínűségi változó várható értéke és szórásnégyzete érdekel. Jegyezzük meg, hogy $E\xi_j=E\xi_1=\frac15$, $\operatorname{Var}\xi_j=\operatorname{Var}\xi_1=\frac15-\left(\frac15\right)^2=\frac{4}{25}$ minden $1\le j\le 10$ indexre, és $\operatorname{Cov}(\xi_j,\xi_k)=E\xi_j\xi_k-E\xi_jE\xi_k=E\xi_1\xi_2-E\xi_1E\xi_2=\frac15\cdot\frac{9}{49}-\left(\frac15\right)^2=-\frac{4}{1225}$ minden $1\le j,k\le 10$, $j\ne k$ indexre. Innen $ES=\sum\limits_{j=1}^{10}E\xi_j=2$, és $\operatorname{Var}S=\sum\limits_{j=1}^{10}\operatorname{Var}\xi_j+2\sum\limits_{j=1}^{9}\sum\limits_{k=j+1}^{10}\operatorname{Cov}(\xi_j,\xi_k)=10\cdot\frac{4}{25}-90\cdot\frac{4}{1225}=\frac{64}{49}$.

2.) Jelölje $G(x)$ a ξ^3 valószínűségi változó eloszlás és $g(x)$ ξ^3 sűrűségfüggvényét. Ekkor $G(x)=P(\xi^3<x)=P(\xi<x^{1/3})=F(x^{1/3})$, ahol $F(x)=1-e^{-\lambda x}$, ha $x\ge 0$, $F(x)=0$, ha $x<0$. Ezért $G(x)=0$, ha $x<0$, $G(x)=1-e^{-\lambda x^{1/3}}$. Mivel $g(x)=\frac{dG(x)}{dx}$, ezért $g(x)=0$, ha $x<0$, $g(x)=\frac13\lambda x^{-2/3}e^{-\lambda x^{1/3}}$, ha $x\ge 0$.

3.) *Első megoldás:* $E\xi^5=\int_0^1 x^5\,dx=\left[\frac{x^6}{6}\right]_0^1=\frac16$, és $\operatorname{Var}\xi^5=E(\xi^5)^2-(E\xi^5)^2=E\xi^{10}-\left(\frac16\right)^2$, $E\xi^{10}=\int_0^1 x^{10}\,dx=\left[\frac{x^{11}}{11}\right]_0^1=\frac{1}{11}$, ahonnan $\operatorname{Var}\xi^5=\frac{1}{11}-\frac{1}{36}=\frac{25}{396}$.

Második megoldás: Számoljuk ki először az $\eta=\xi^5$ valószínűségi változó $g(x)$ sűrűségfüggvényét. $G(x)=P(\eta<x)=P(\xi<x^{1/5})$, $G(x)=x^{1/5}$, ha $0\le x\le 1$, és $G(x)=0$, ha $x<0$, $G(x)=1$, ha $x>1$. Innen $g(x)=\frac{dG(x)}{dx}$, $g(x)=\frac15 x^{-4/5}$, ha $0\le x\le 1$, $g(x)=0$, ha $x<0$ vagy $x>1$. Ezért $E\xi^5=E\eta=\int_0^1 x\cdot\frac15 x^{-4/5}\,dx=\frac15\left[\frac56 x^{6/5}\right]_0^1=\frac16$, $\operatorname{Var}\xi^5=\operatorname{Var}\eta=E\eta^2-(E\eta)^2=\int_0^1 x^2\cdot\frac15 x^{-4/5}\,dx-\left(\frac16\right)^2=\frac15\left[\frac{5}{11}x^{11/5}\right]_0^1=\frac{1}{11}-\frac{1}{36}=\frac{25}{396}$.

4.) Tudjuk, hogy

$$E\xi=\int_0^{2\pi} x\frac{1}{2\pi}\left(1+\frac12\sin x\right)=\frac{1}{2\pi}\int_0^{2\pi}x\,dx+\frac{1}{4\pi}\int_0^{2\pi}x\sin x\,dx,$$

$\mathrm{Var}\,\xi = E\xi^2 - (E\xi)^2$, és

$$E\xi^2 = \int_0^{2\pi} x^2 \frac{1}{2\pi}\left(1+\frac{1}{2}\sin x\right) = \frac{1}{2\pi}\int_0^{2\pi} x^2\,dx + \frac{1}{4\pi}\int_0^{2\pi} x^2 \sin x\,dx.$$

Továbbá $\frac{1}{2\pi}\int_0^{2\pi} x\,dx = \frac{1}{2\pi}\frac{(2\pi)^2}{2} = \pi$, és $\frac{1}{2\pi}\int_0^{2\pi} x^2\,dx = \frac{1}{2\pi}\frac{(2\pi)^3}{3} = \frac{4}{3}\pi^2$,. Parciális integrálással ($x\sin x = -x\frac{d}{dx}\cos x$ szereposztással) kapjuk, hogy

$$\begin{aligned}\frac{1}{4\pi}\int_0^{2\pi} x\sin x\,dx &= \frac{1}{4\pi}\left[-x\cos x\right]_0^{2\pi} + \frac{1}{4\pi}\int_0^{2\pi}\cos x\,dx \\ &= \frac{1}{4\pi}\left(-2\pi + [\sin x]_0^{2\pi}\right) = -\frac{1}{2}.\end{aligned}$$

Hasonlóan, kétszeres parciális integrálás adja, hogy

$$\begin{aligned}\frac{1}{4\pi}\int_0^{2\pi} x^2\sin x\,dx &= \frac{1}{4\pi}\left[-x^2\cos x\right]_0^{2\pi} + \frac{1}{4\pi}\int_0^{2\pi} 2x\cos x\,dx \\ &= -\pi + \frac{1}{4\pi}\left[2x\sin x\right]_0^{2\pi} - \frac{1}{2\pi}\int_0^{2\pi}\sin x\,dx = -\pi + \frac{1}{4\pi}\left[2x\sin x\right]_0^{2\pi} = -\pi.\end{aligned}$$

Innen $E\xi = \pi - \frac{1}{2}$, $E\xi^2 = \frac{4}{3}\pi^2 - \pi$, $\mathrm{Var}\,\xi = \frac{4}{3}\pi^2 - \pi - (\pi - \frac{1}{2})^2 = \frac{\pi^2}{3} - \frac{1}{4}$.

5.) Vezessük be a következő ξ_j, $1 \le j \le 1200$, valószínűségi változókat: $\xi_j = 2$, ha a J-ik dobás értéke 2, $\xi_j = 4$, ha a J-ik dobás értéke 4, $\xi_j = 6$, ha a J-ik dobás értéke 6, $\xi_j = 0$, ha a J-ik dobás értéke 1, 3 vagy 5. Ekkor a ξ_j valószínűségi változók függetlenek és egyforma eloszlásúak, és az $S = \sum\limits_{j=1}^{1200} \xi_j$ jelöléssel minket a $P(2280 < S < 2500)$ valószínűség érdekel. Továbbá, $E\xi_j = \frac{1}{6}(2+4+6) = 2$, $E\xi_j^2 = \frac{1}{6}(2^2+4^2+6^2) = \frac{28}{3}$, $\mathrm{Var}\,\xi_j = E\xi_j^2 - (E\xi_j)^2 = \frac{1}{6}3$, ahonnan $ES = 2400$, $\mathrm{Var}\,S = 1200 \cdot \frac{16}{3} = (80)^2$. Innen $P(2280 < S < 2500) = P(-1.5 < \frac{S-ES}{\sqrt{S}} < 1.25)$, és a centrális határeloszlástétel szerint $P(-1.5 < \frac{S-ES}{\sqrt{S}} < 1.25) \sim \Phi(1.25) - \Phi(-1.5) = \Phi(1.5) + \Phi(1.25) - 1 \sim 0.8944 + 0.933 - 1 = 0.8276$. A keresett valószínűség tehát körülbelül 0.83.

6.) **Valószínűségi változók függetlenségének a definiciója.** *Legyenek* $\xi_1, \dots, \xi_n$ *valószínűségi változók egy* $(\Omega, \mathcal{A}, P)$ *valószínűségi mezőn. Azt mondjuk, hogy ezek a valószínűségi változók függetlenek, ha minden* $x_1, \dots, x_n$ *valós számra*

$$P(\xi_1 < x_1, \dots, \xi_n < x_n) = P(\xi_1 < x_1) \cdots P(\xi_n < x_n).$$

Elfogadtam a következő választ is, mert a valószínűségszámítás bizonyos eredményei szerint ez ekvivalens az eredeti definicióval.

A $\xi_1, \dots, \xi_n$ *valószínűségi változók akkor függetlenek, ha a számegyenes minden* $B_1, \dots, B_n$ *Borel mérhető részhalmazára teljesül a*

$$P(\xi_1 \in B_1, \dots, \xi_n \in B_n) = P(\xi_1 \in B_1) \cdots P(\xi_n \in B_n)$$

azonosság.